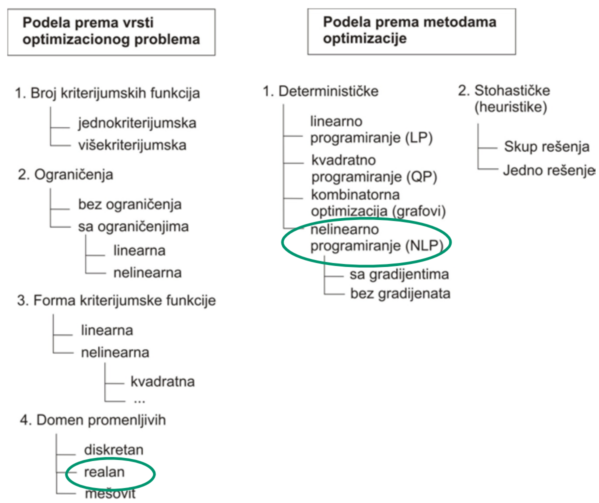
METODE OPTIMIZACIJE
NELINEARNO PROGRAMIRANJE

Dr Tina Dašić
Dr Miloš Stanić

Građevinski fakultet Univerziteta u Beogradu
2014.

UVOD



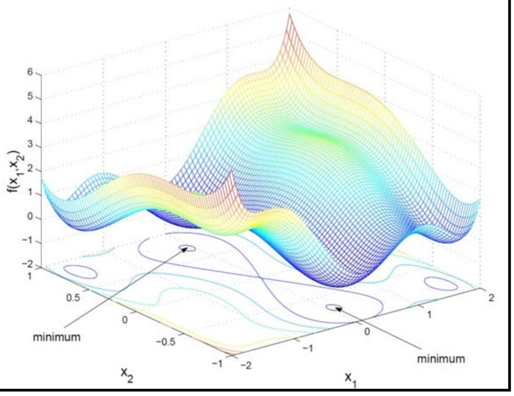
FORMULACIJA PROBLEMA

Zadatak optimizacije je pronalaženje promenljivih pri kojima ciljna (kriterijumska) funkcija uzima ekstremnu (minimalnu ili maksimalnu) vrednost, uz ograničenja sa kojima se definiše prostor mogućih rešenja.

$$\min_{X \in R^n} f(X)$$

$$\varphi_j(X) = 0,\ j = 1,\ldots,J$$

$$\psi_k(X) \le 0,\ k = 1,\ldots,K$$



f kriterijumska funkcija,
φ ograničenja u formi jednakosti
Ψ funkcije ograničenja u formi nejednakosti.

FORMULACIJA PROBLEMA

Ako se radi o problemu određivanja maksimuma funkcije, on se dobija veoma jednostavnom reformulacijom problema:

$$\max f(X) = \min (-f(X))$$

Ako su kriterijumska funkcija i funkcije ograničenja linearne, radi se o zadatku linearnog programiranja.

Ukoliko funkcije ograničenja ne postoje, onda se radi o problemu bezuslovne optimizacije.

$$\min_{X \in R^n} f(X) \longrightarrow X^* = \begin{bmatrix} x_1 \\ \ldots \\ x_n \end{bmatrix}$$

FORMULACIJA PROBLEMA

Ukoliko funkcije ograničenja postoje, moguće je preformulisati optimizacioni zadatak uvođenjem kaznene funkcije *P(X)*.

$$\begin{gathered}\min_{X\in R^n} f(X)\\ \varphi_j(X)=0,\ j=1,\dots,J\\ \psi_k(X)\le 0,\ k=1,\dots,K\end{gathered} \longrightarrow \begin{gathered}\min_{X\in R^n} f^*(X)=f(X)+\lambda\cdot P(X)\\ P(X)=\sum_j\left[\varphi_j(X)\right]^2+\sum_k\left[\max(0,\psi_k(X))\right]^2\end{gathered}$$

Očigledno je da je vrednost kaznene funkcije *P(X)=0*, kada su sva ograničenja zadovoljena, čime je moguće optimizacioni zadatak sa ograničenjima svesti na zadatak bez ograničenja.

Bezuslovna optimizacija - funkcija jedne promenljive

Uslovi optimalnosti

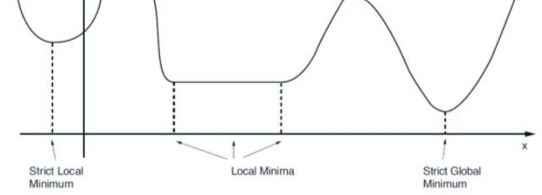
Egzistencija rešenja - da li rešenje postoji?

Potreban uslov: $f'(x)=0$

Dovoljan uslov:
ako je x^* tačka u kojoj je ispunjen potreban uslov pojave lokalnog ekstrema $f'(x^*)=0$ i ukoliko je $f''(x^*)>0$, onda je x^* tačka lokalnog minimuma.

Razvojem funkcije ***f(x)*** u Taylor-ov red u okolini x^*:

$$f(x)=f(x^*)+\overset{0}{f'(x^*)}(x-x^*)+\frac{1}{2}f''(x^*)(x-x^*)^2+\dots$$



Bezuslovna optimizacija - funkcija jedne promenljive

Metode zasnovane na gradijentima – Newton-ova metoda

Razvojem funkcije ***f(x)*** u Taylor-ov red dobija se:

$$f(x) = f(x_i) + f'(x_i)(x - x_i) + \frac{1}{2} f''(x_i)(x - x_i)^2 + ...$$

Postavljanjem potrebnog uslova *f'(x)=0,* uz zadržavanje prva 3 člana reda, dobija se:

$$f'(x_i) + f''(x_i)(x - x_i) = 0$$

Imajući u vidu da su zadržana samo prva tri člana Tejlorovog reda, do rešenja se stiže u iteracijama:

$$x_{i+1} = x_i - \frac{f'(x_i)}{f''(x_i)}$$

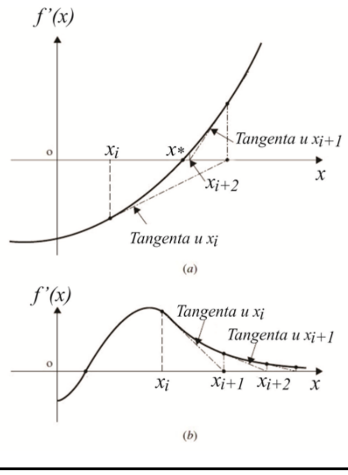
Bezuslovna optimizacija - funkcija jedne promenljive

Metode zasnovane na gradijentima – Newton-ova metoda

$$x_{i+1} = x_i - \frac{f'(x_i)}{f''(x_i)}$$

Problem se svodi na pronalaženje nule funkcije *f' (x)*.

Uslovi za uspešnu primenu metode:
• poznati prvi i drugi izvod funkcije
• rezultat može zavisiti od izbora tačke od koje se počinje sa pretragom (slika a) – primer uspešne primene metode, a slika b) primer neuspešne)



Bezuslovna optimizacija - funkcija jedne promenljive

Metode zasnovane na gradijentima – Newton-ova metoda

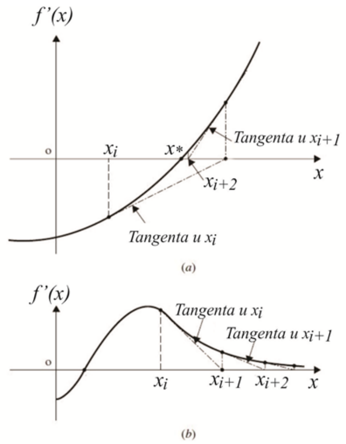
$$x_{i+1} = x_i - \frac{f'(x_i)}{f''(x_i)}$$

Ako prvi ili drugi izvod funkcije nisu poznati ili ih je teško odrediti, $f'(x_i)$ i $f''(x_i)$ se mogu sračunati približno, primenom metode konačnih razlika:

$$f'(x_i) = \frac{f(x_i+\delta) - f(x_i-\delta)}{2\delta}$$

$$f''(x_i) = \frac{f(x_i+\delta) - 2f(x_i) + f(x_i-\delta)}{\delta^2}$$

$$x_{i+1} = x_i - \frac{\delta \cdot (f(x_i+\delta) - f(x_i-\delta))}{2 \cdot (f(x_i+\delta) - 2f(x_i) + f(x_i-\delta))}$$



Bezuslovna optimizacija - funkcija više promenljivih

Uslovi optimalnosti

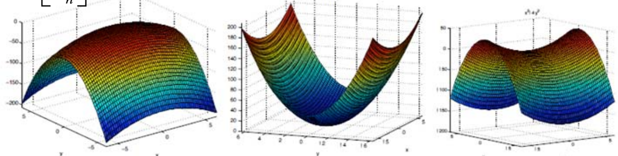
Egzistencija rešenja - da li rešenje postoji?

$$\min_{X \in R^n} f(X) \longrightarrow X^* = \begin{bmatrix} x_1 \\ \dots \\ x_n \end{bmatrix}$$

Gradijent funkcije: $\nabla f(X) = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \dots \\ \frac{\partial f}{\partial x_n} \end{bmatrix}$

Potreban uslov lokalnog ekstrema:

$$\nabla f(X) = 0 \longrightarrow X^* = \begin{bmatrix} x_1 \\ \dots \\ x_n \end{bmatrix}$$



Maximum Minimum Saddle point

Bezuslovna optimizacija - funkcija više promenljivih

Uslovi optimalnosti

Potreban uslov lokalnog ekstrema:

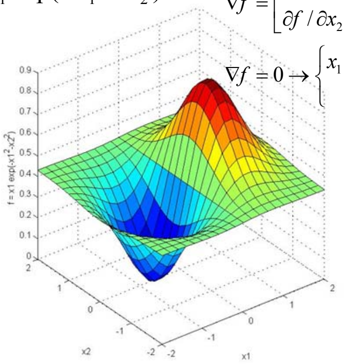
$$\nabla f(X) = 0 \longrightarrow X^* = \begin{bmatrix} x_1 \\ \dots \\ x_n \end{bmatrix}$$

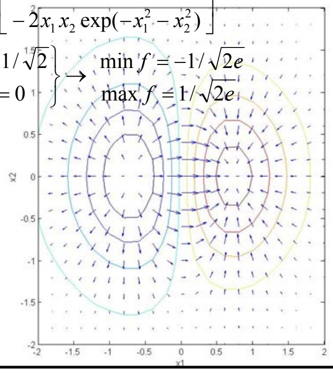
Primer

$$f = x_1 \exp(-x_1^2 - x_2^2)$$

$$\nabla f = \begin{bmatrix} \partial f / \partial x_1 \\ \partial f / \partial x_2 \end{bmatrix} = \begin{bmatrix} (1-2x_1^2)\exp(-x_1^2 - x_2^2) \\ -2x_1 x_2 \exp(-x_1^2 - x_2^2) \end{bmatrix}$$

$$\nabla f = 0 \rightarrow \left\{ \begin{matrix} x_1 = \pm 1/\sqrt{2} \\ x_2 = 0 \end{matrix} \right\} \rightarrow \begin{matrix} \min f = -1/\sqrt{2e} \\ \max f = 1/\sqrt{2e} \end{matrix}$$



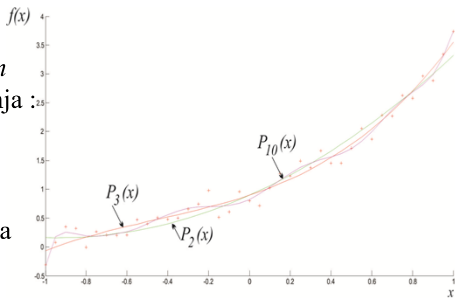


Bezuslovna optimizacija - funkcija više promenljivih

Uslovi optimalnosti

Potreban uslov lokalnog ekstrema: $\nabla f(X) = 0$

Primer – metod najmanjih kvadrata



Problem pronalaženja polinoma stepena n (P_n) sa kojim treba „fitovati“ tačke merenja : $\{(x_1,y_1), \dots, (x_m,y_m)\}$

$$P_n(x) = a_0 + a_1 x + \dots + a_n x^n$$

Potrebno je odrediti koeficijente polinoma (vektor) $a=[a_0,\dots,a_n]^T$, koji predstavljaju minimum funkcije greške:

$$f(a) = \sum_{i=1}^{m} (y_i - P_n(x_i))^2 \longrightarrow \nabla f = \begin{bmatrix} \partial f / \partial a_0 \\ \dots \\ \partial f / \partial a_n \end{bmatrix} = 0$$

Dobija se sistem linearnih jednačina koji se rešava po nepoznatom vektoru a.

Bezuslovna optimizacija - funkcija više promenljivih

Primer – metod najmanjih kvadrata

Ovaj problem se može napisati i u matričnom obliku:

$$b = \begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_m \end{bmatrix}, \quad A = \begin{bmatrix} 1 & x_1 & \dots & x_1^n \\ 1 & x_2 & \dots & x_2^n \\ \dots & \dots & \dots & \dots \\ 1 & x_m & \dots & x_m^n \end{bmatrix}, \quad a = \begin{bmatrix} a_0 \\ a_1 \\ \dots \\ a_n \end{bmatrix} \qquad r = A \cdot a - b = \begin{bmatrix} P_n(x_1) - y_1 \\ \dots \\ P_n(x_m) - y_m \end{bmatrix}$$

gde je sa r označen vektor reziduala. Funkcija, koja predstavlja sumu kvadrata odstupanja ima oblik:

$$f(a) = r^T r = a^T A^T A a - 2b^T A a + b^T b$$

Iz potrebnog uslova optimalnosti , dobija se:

$$\nabla f(a) = 2A^T A a - 2A^T b = 0 \longrightarrow \boxed{a = (A^T A)^{-1} A^T b}$$

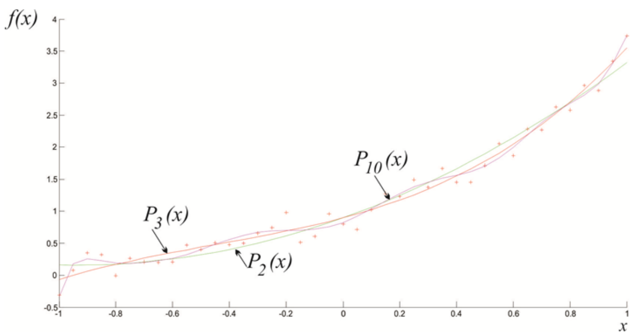
Bezuslovna optimizacija - funkcija više promenljivih

Primer – metod najmanjih kvadrata

Primer:
m = 40

Korišćeni polinomi:
n = 2, 3 i 10

Srednje kvadratno odstupanje:

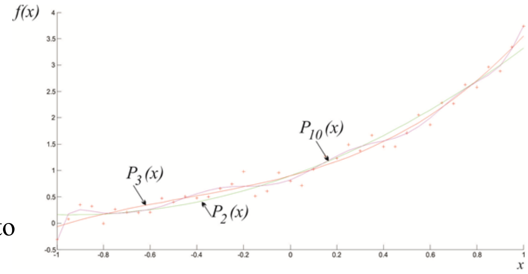


$err = (r^T r)/\mathrm{m} = 0.0415,\ 0.0240$ i 0.0214

Bezuslovna optimizacija - funkcija više promenljivih

Primer – metod najmanjih kvadrata

Sa polinomom većeg stepena, moguće je dobiti bolje slaganje, ali to podrazumeva pojavu “oscilacija” koje najčešće nisu deo prirode fenomena.



Zbog toga se funkcija greške često predstavlja na sledeći način:

$$f(a) = r^T r + \gamma\, a^T a$$

gde je γ (>0) koeficijent kojim se penalizuju polinomi višeg stepena.

$$a = (A^T A + \gamma\, E)^{-1} A^T b$$

E – jedinična matrica dimenzija ($n+1$)

Bezuslovna optimizacija - funkcija više promenljivih

Uslovi optimalnosti

Potreban uslov lokalnog ekstrema: $\nabla f(X) = 0 \longrightarrow X^* = \begin{Bmatrix} x_1 \\ \dots \\ x_n \end{Bmatrix}$

<u>Dovoljan uslov minimuma *f(X)*:</u>

Razvojem funkcije *f(X)* u Taylor-ov red u okolini tačke X^* i zadržavanjem samo prva tri člana: $f(X) = f(X^*) + (\nabla f(X^*))^T (X - X^*) + \frac{1}{2}(X - X^*)^T (\nabla^2 f(X^*))(X - X^*)$

Matrica Hessian-a: $H(X) = \nabla^2 f(X) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \dots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \dots & \dots & \dots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \dots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$

Ako je X^* tačka u kojoj je zadovoljen potreban uslov optimalnosti $\nabla f(X^*) = 0$ da bi ta tačka bila minimum, potrebno je da $f(X) > f(X^*)$ iz čega sledi dovoljan uslov minimuma funkcije više promenljivih:

$$\boxed{\Delta X^T H(X^*)\, \Delta X > 0}$$

gde je: $\Delta X = X - X^*$

Bezuslovna optimizacija - funkcija više promenljivih

Uslovi optimalnosti

Potreban uslov lokalnog ekstrema: $\boxed{\nabla f(X)=0} \longrightarrow X^* = \begin{Bmatrix} x_1 \\ \dots \\ x_n \end{Bmatrix}$

Dovoljan uslov minimuma funkcije više promenljivih, je da matrica Hessian-a ***H*** bude *pozitivno definitna*:

$$\boxed{\Delta X^T H(X^*)\,\Delta X > 0}$$

Po definiciji, pozitivna definitnost matrice podrazumeva da prethodni uslov bude ispunjen za bilo koji vektor ΔX.

Kako prethodni uslov mora biti ispunjen i za svaki jedinični vektor, to znači da dijagonalni elementi matrice H moraju biti pozitivni (nije dovoljan uslov).

Postoji nekoliko načina da se proveri pozitivna definitnost matrice H.

Bezuslovna optimizacija - funkcija više promenljivih

Uslovi optimalnosti

Simetrična matrica je pozitivno definitna, ako su svi minori glavne dijagonale H_{ii} pozitivni.

Primer: $f = 4x_1^2 + 3x_2^2 + 5x_3^2 + 6x_1x_2 + x_1x_3 - 3x_1 - 2x_2 + 15$:

$$\mathbf{H(X)} = \begin{bmatrix} \partial^2 f/\partial x_1^2 & \partial^2 f/\partial x_1\,\partial x_2 & \partial^2 f/\partial x_1\,\partial x_3 \\ \partial^2 f/\partial x_1\,\partial x_2 & \partial^2 f/\partial x_2^2 & \partial^2 f/\partial x_2\,\partial x_3 \\ \partial^2 f/\partial x_1 \partial x_3 & \partial^2 f/\partial x_2\,\partial x_3 & \partial^2 f/\partial x_3^2 \end{bmatrix} = \begin{bmatrix} 8 & 6 & 1 \\ 6 & 6 & 0 \\ 1 & 0 & 10 \end{bmatrix}$$

Determinante glavnih minora: H_{11}, H_{22}, H_{33} su pozitivne, pa je matrica pozitivno definitna:

$$|H_{11}| = |8| = 8 > 0$$

$$|H_{22}| = \begin{vmatrix} 8 & 6 \\ 6 & 6 \end{vmatrix} = 8 > 0$$

$$|H_{33}| = \begin{vmatrix} 8 & 6 & 1 \\ 6 & 6 & 0 \\ 1 & 0 & 10 \end{vmatrix} = 114 > 0$$

Bezuslovna optimizacija - funkcija više promenljivih

Uslovi optimalnosti

Drugi način za proveru pozitivne definitnosti matrice je preko sopstvenih vrednosti.

$$HX = \lambda X$$

$$(H - \lambda E)X = 0$$

Sopstvene vrednosti λ se dobijaju iz uslova da netrivijalna rešenja sistema postoje kada je ispunjen uslov:

$$\left| H - \lambda E \right| = 0 \rightarrow \lambda$$

Ako su sve sopstvene vrednosti, simterične matrice H pozitivne, matrica je pozitivno definitna: $X^T HX = X^T \lambda X \rightarrow \lambda = \dfrac{X^T HX}{X^T X} > 0$

Primer:

$$H = \begin{bmatrix} \alpha & \beta \\ \beta & \alpha \end{bmatrix}$$

$$\left| H - \lambda E \right| = \begin{vmatrix} \alpha - \lambda & \beta \\ \beta & \alpha - \lambda \end{vmatrix} = (\alpha - \lambda)^2 - \beta^2 = 0 \longrightarrow \lambda = \alpha \pm \beta > 0$$

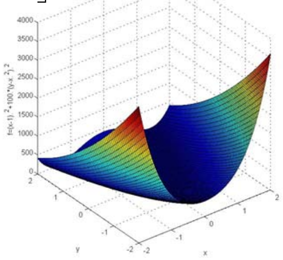
Bezuslovna optimizacija - funkcija više promenljivih

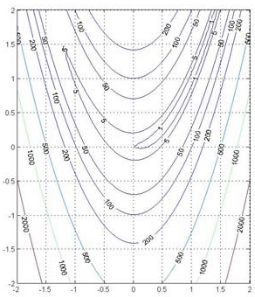
Uslovi optimalnosti

Primer: *"Banana funkcija"* $f(x, y) = (x-1)^2 + 100(y - x^2)^2$

Potreban uslov:

$$\nabla f = \begin{bmatrix} 2(x-1) - 400x(y - x^2) \\ 200(y - x^2) \end{bmatrix} = 0 \rightarrow x^* = 1, \, y^* = 1$$





Dovoljan uslov:

$$H = \begin{bmatrix} 2 + 1200x^2 - 400y & -400x \\ -400x & 200 \end{bmatrix} \rightarrow H(1,1) = \begin{bmatrix} 802 & -400 \\ -400 & 200 \end{bmatrix}$$

Bezuslovna optimizacija - funkcija više promenljivih

Metode zasnovane na gradijentima

Metoda Newton-a

$$f(X) = f(X_i) + (\nabla f(X_i))^T (X - X_i) + \frac{1}{2}(X - X_i)^T (\nabla^2 f(X_i))(X - X_i)$$

Iz uslova da je prvi izvod funkcije *f(X)* jednak 0 dobija se:

$$\nabla f(X_i) + (\nabla^2 f(X_i))(X - X_i) = 0$$

Do rešenja se dolazi u iteracijama:

$$X_{i+1} = X_i - H(X_i)^{-1} \nabla f(X_i)$$

Prednost ove metode je brzina konvergencije. Kod kvadratnih funkcija, rešenje se dobija u jednom koraku.
Nedostatak je, što je potrebno računati matricu *H* i njenu inverznu vrednost u svakoj iteraciji proračuna. Osim toga proračun je veoma osetljiv na izbor početne tačke.

Bezuslovna optimizacija - funkcija više promenljivih

Metode zasnovane na gradijentima

Metoda Newton-a

Primer:

$$f = (x_1 - 3)^2 + 3(x_2 - 1)^2 + 2$$

Originalna Newton-ova metoda zahteva poznavanje prvog i drugog izvoda funkcije:

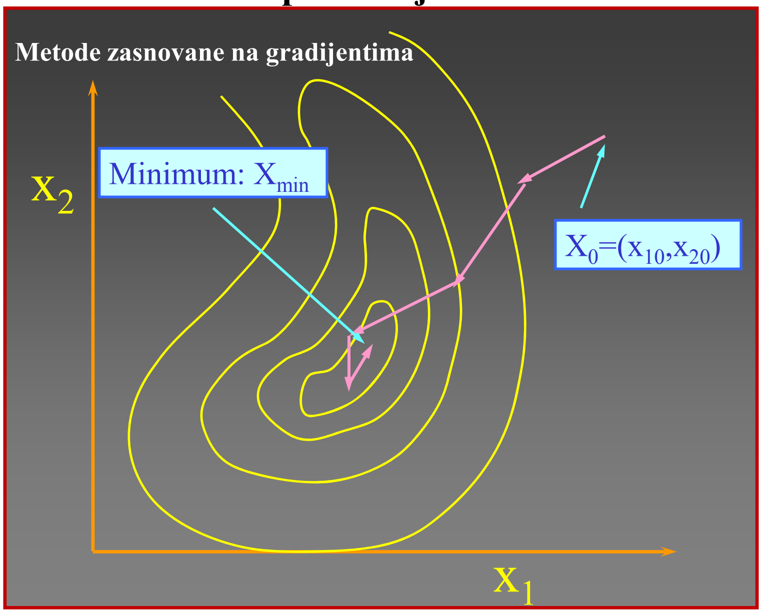
$$\nabla f = \begin{bmatrix} 2(x_1 - 3) \\ 6(x_2 - 1) \end{bmatrix} \qquad H = \begin{bmatrix} 2 & 0 \\ 0 & 6 \end{bmatrix}$$

Usvaja se početna tačka: $X_0=[0, 0]^T$

$$X_1 = X_0 - H^{-1} \nabla f(X_0) = \begin{bmatrix} 0 \\ 0 \end{bmatrix} - \begin{bmatrix} 1/2 & 0 \\ 0 & 1/6 \end{bmatrix} \cdot \begin{bmatrix} -6 \\ -6 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

U tački $X_1=[3, 1]^T$, ispunjeni su i potreban i dovoljan uslov, pa je ta tačka lokalni minimum.

Bezuslovna optimizacija - funkcija više promenljivih



Bezuslovna optimizacija - funkcija više promenljivih

Metode zasnovane na gradijentima

Metod najstrmijih gradijenata – Cauchy-jev metod

Kod metode najstrmijih gradijenata, pronalazi se najmanja vrednost funkcije u pravcu vektora *Δs*.

$$f(X_i+\Delta s)=f(X_i)+(\nabla f(X_i))^T\Delta s$$

Kako je ideja da se nadje manja vrednost funkcije to mora biti ispunjen uslov:

$$f(X_i+\Delta s)-f(X_i)=(\nabla f(X_i))^T\Delta s<0$$

Scalarni proizvod dva vektora je minimalan kada su oni paralelni i suprotnim smerovima, pa *Δs* treba računati iz uslova: $\Delta s=-\alpha\nabla f(X_i)$

gde je α > 0 veličina koraka propagacije u pravcu *Δs*.

Bezuslovna optimizacija - funkcija više promenljivih

Metode zasnovane na gradijentima

Metod najstrmijih gradijenata – Cauchy-jev metod

$$f(X_{i+1}) = f(X_i) - \alpha(\nabla f(X_i))^T \nabla f(X_i)$$

Vrednost α, nije proizvoljna. Mala vrednost, bi značila sporu konvergenciju, a prevelika bi mogla dovesti da se promaši cilj i ode dalje od lokalnog minimuma.

Korak α se menja u svakoj iteraciji proračuna i dobija se iz uslova minimumu funkcije:

$$\phi(\alpha_i) = f(X_{i+1}) = f(X_i - \alpha_i \nabla f(X_i))$$

$$\frac{d\phi(\alpha_i)}{d\alpha_i} = 0 \to \alpha_i$$

Nova vrednost vektora: $X_{i+1} = X_i - \alpha_i \nabla f(X_i)$

Ako se Cauchy-jev metod uporedi sa metodom Newton-a, očigledno je da je matrica *H* pojednostavljena i da se računa: $H(X_i)^{-1} = \alpha_i E$

Bezuslovna optimizacija - funkcija više promenljivih

Metode zasnovane na gradijentima

Metod najstrmijih gradijenata – Cauchy-jev metod

Primer: $f = 10x_1^2 + 5x_1x_2 + 10(x_2 - 3)^2$, početna tačka je X_0=[10, 15]^T

Prva iteracija: $X_1 = X_0 - \alpha_0 \nabla f(X_0) = \begin{bmatrix} 10 \\ 15 \end{bmatrix} - \alpha_0 \begin{bmatrix} 275 \\ 290 \end{bmatrix}$

Korak propagacije α_0 se nalazi iz uslova minimuma funkcije *f* u pravcu gradijenta u tački X_0:

$$\phi(\alpha_0) = 10(10 - 275\alpha_0)^2 + 5(10 - 275\alpha_0)(15 - 290\alpha_0) + 10(12 - 290\alpha_0)^2$$

$$\frac{d\phi(\alpha_0)}{d\alpha_0} = -159725 + 3992000\alpha_0 = 0 \to \alpha_0 = 0.04001 \longrightarrow X_1 = [-1.0027, 3.3971]^T$$

Druga iteracija: $\nabla f(X_1) = \begin{bmatrix} -3.078 \\ 2.919 \end{bmatrix}$, $X_2 = X_1 - \alpha_1 \begin{bmatrix} -3.078 \\ 2.919 \end{bmatrix}$

Iz minimuma funkcije $\phi(\alpha_1)$, dobija se α_1=0.066, pa je X_2=[-0.797,3.202]^T

Treća iteracija:

$$\nabla f(X_2) = \begin{bmatrix} 0.060 \\ 0.064 \end{bmatrix}, \quad X_3 = X_2 - \alpha_2 \begin{bmatrix} 0.060 \\ 0.064 \end{bmatrix} \to \alpha_2 = 0.040 \to X_3 = \begin{bmatrix} -0.8000299 \\ 3.20029 \end{bmatrix}$$

Dobijeno rešenje je veoma blisko tačnom: X^* = [-4/5,16/5]^T = [-0.8,3.2]^T

Bezuslovna optimizacija - funkcija više promenljivih

Metode zasnovane na gradijentima

Metod najstrmijih gradijenata – Cauchy-jev metod

Nedostatak Cauchy-jeve metode je što se u svakoj iteraciji traži tačna veličina koraka u pravcu najstrmijeg pada, što može biti veoma složeno. Za rešavanje realnih problema, najčešće tačna vrednost koraka α nije tako važna, već bi pronalaženje približne vrednosti bilo dovoljno dobro.

Ako se α, traži na približan način u svakoj iteraciji proračuna, moraju biti ispunjeni određeni uslovi, kako izabrana vrednost ne bi bila prevelika ili premala.

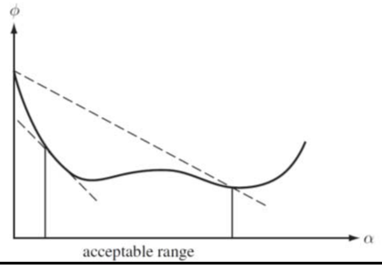
Wolfe-ovi uslovi postavljaju ograničenja po pitanju vrednosti funkcije i prvog izvoda

$$\phi(\alpha) \le \phi(0) + \varepsilon \cdot \phi'(0) \cdot \alpha$$

$$\phi'(\alpha) > (1-\varepsilon) \cdot \phi'(0)$$

Gde je ε parametar koji uzima vrednost 0 do ½.

Algoritam za pronalaženje prihvatljive vrednosti α: krene se sa maksimalnom vrednošću i ispita se ispunjenost Wolfe-ovih uslova. Ako uslovi nisu ispunjeni, korak se redukuje množenjem sa faktorom koji je manji od 1 (npr. 0.9). Postupak se ponavlja dok uslovi nisu ispunjeni.



Bezuslovna optimizacija - funkcija više promenljivih

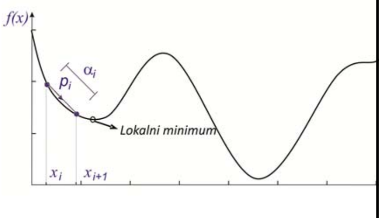
Metode zasnovane na gradijentima

Osnovni koraci optimizacionih metoda zasnovanih na gradijentima:

Definiše se početna tačka X_0

$i = 0,1,2,.....$

- proveriti konvergenciju u tački X_i
- sračunati pravac pomeranja p_i
- odrediti veličinu koraka pomeranja α_i
- odrediti položaj nove tačke $X_{i+1} = X_i + \alpha_i p_i$



Sve metode zasnovane na gradijentima se razlikuju po tome na koji način se računaju pravac pomeranja i veličina koraka.

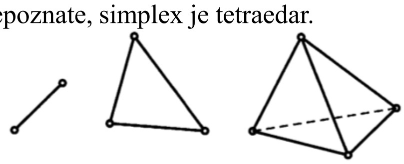
Bezuslovna optimizacija - funkcija više promenljivih

Metode bez gradijenata

Kada izvode funkcije nije moguće odrediti eksplicitno, jedna mogućnost je da se to uradi približno - numerički, a druga je da se proračun sprovede primenom neke metode koja ne zahteva poznavanje izvoda funkcije.

Metoda Nelder-Mead-a:

Kod ove metode, za proračun se koristi skup rešenja, koji se pomera iz jedne iteracije u drugu. Brojnost ovog skupa je za jedan veći od broja nepoznatih, odnosno od dimenzija vektora *X*.
Iterativni skup *P,* koji ima *n*+1 tačku u *n* dimenzionalnom prostoru, se naziva simplex. Kada je problem dvodimenzionalan, simpleks je trougao. U slučaju funkcije sa 3 nepoznate, simplex je tetraedar.



(a) (b) (c)

Bezuslovna optimizacija - funkcija više promenljivih

Metode bez gradijenata

Metoda Nelder-Mead-a

<u>Osnovni koraci</u>

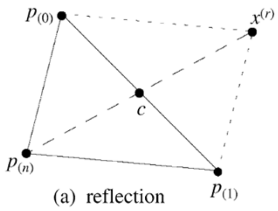
Prvi korak je da se formira skup *P* koji se zatim sortira prema vrednosti kriterijumske funkcije $f(X_i)$:

$$P = \{X_0, \ldots, X_n)$$
$$f(X_0) < f(X_1) < \ldots < f(X_n)$$

Zatim se nalazi centar simpleksa ali se prethodno izostavi tačka n+1, koja je najlošija tačka iz skupa *P*:

$$X_c = \frac{1}{n}\sum_{0}^{n-1} X_i$$

Na osnovu tačke X_c nalazi se tačka X_r, koja predstavlja refleksiju najlošije tačke skupa *P*: $X_r = X_c + \alpha(X_c - X_{n+1})$, gde je α, koeficijent veći od 0 (obično 1).



(a) reflection

Bezuslovna optimizacija - funkcija više promenljivih

Metode bez gradijenata

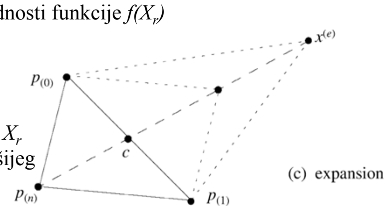
Metoda Nelder-Mead-a

<u>Osnovni koraci</u>

Da li se nova tačka prihvata ili ne, zavisi od vrednosti funkcije $f(X_r)$:

• Ako je $f(X_0) < f(X_r) < f(X_{n-1})$, onda se najlošije rešenje X_n, izbacuje iz skupa P a uvodi se novo rešenje X_r, u sortirani skup P, na mesto koje odgovara vrednosti funkcije $f(X_r)$

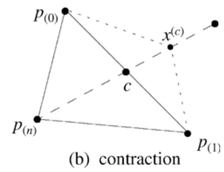
• Ako je $f(X_r) < f(X_0)$, proba se sa daljom ekspanzijom simpleksa u tom pravcu, radi provere da li se vrednost funkcije može još više poboljšati: $X_e = X_r + \alpha(X_r - X_c)$
Ako je $f(X_e) < f(X_r)$ prihvata se rešenje X_e, u protivnom se X_r koristi kao novo rešenje i uvodi se u skup P umesto najlošijeg rešenja X_n.



(c) expansion

• Ako nema napredka $f(X_r) > f(X_{n-1})$, kontrakuje se simpleks tako što se najlošija tačka približava centru: $X_k = X_n + \gamma(X_c - X_n)$

Gde γ uzima vrednost 0.5. Ako $f(X_k) < f(X_n)$, X_k se koristi kao novo rešenje i uvodi se u skup P umesto najlošijeg rešenja X_n.



(b) contraction

Bezuslovna optimizacija - funkcija više promenljivih

Metode bez gradijenata

Metoda Nelder-Mead-a

<u>Osnovni koraci</u>

Da li se nova tačka prihvata ili ne, zavisi od vrednosti funkcije $f(X_r)$:

• Ako nijedan od navedenih uslova nije ispunjen, ceo skup P se kontrakuje $X_i = X_1 + \gamma(X_i - X_1)$ prema najboljem rešenju X_1: